

## Online information on face masks: analysis of websites in Italian and English returned by different search engines

Article (Published Version)

Mehta, Shaily, Ghezzi, Daria, Catalani, Alessia, Vanzolini, Tania and Ghezzi, Pietro (2021) Online information on face masks: analysis of websites in Italian and English returned by different search engines. *BMJ Open*, 11. e046364 1-8. ISSN 2044-6055

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/101630/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

### **Copyright and reuse:**

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

# BMJ Open Online information on face masks: analysis of websites in Italian and English returned by different search engines

Shaily Mehta,<sup>1</sup> Daria Ghezzi,<sup>2</sup> Alessia Catalani,<sup>3</sup> Tania Vanzolini,<sup>3</sup> Pietro Ghezzi <sup>1</sup>

**To cite:** Mehta S, Ghezzi D, Catalani A, *et al.* Online information on face masks: analysis of websites in Italian and English returned by different search engines. *BMJ Open* 2021;**11**:e046364. doi:10.1136/bmjopen-2020-046364

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-046364>).

Received 28 October 2020  
Accepted 21 June 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>Brighton and Sussex Medical School, Brighton, UK

<sup>2</sup>Homerton College, University of Cambridge, Cambridge, UK

<sup>3</sup>School of Biotechnology, University of Urbino Carlo Bo, Urbino, Italy

## Correspondence to

Dr Pietro Ghezzi;  
[p.ghezzi@bsms.ac.uk](mailto:p.ghezzi@bsms.ac.uk)

## ABSTRACT

**Objective** Countries have major differences in the acceptance of face mask use for the prevention of COVID-19. This work aims at studying the information online in different countries in terms of information quality and content.

**Design** Content analysis.

**Method** We analysed 450 webpages returned by searching the string 'are face masks dangerous' in Italy, the UK and the USA using three search engines (Bing, Duckduckgo and Google) in August 2020. The type of website and the stance about masks were assessed by two raters for each language and inter-rater agreement reported as Cohen's kappa. The text of the webpages was collected from the web using WebBootCaT and analysed using a corpus analysis software to identify issues mentioned.

**Results** Most pages were news outlets, and few (2%–6%) from public health agencies. Webpages with a negative stance on masks were more frequent in Italian (28%) than English (19%). Google returned the highest number of mask-positive pages and Duckduckgo the lowest. Google also returned the lowest number of pages mentioning conspiracy theories and Duckduckgo the highest. Webpages in Italian scored lower than those in English in transparency (reporting authors, their credentials and backing the information with references). When issues about the use of face masks were analysed, mask effectiveness was the most discussed followed by hypercapnia (accumulation of carbon dioxide), contraindication in respiratory disease and hypoxia, with issues related to their contraindications in mental health conditions and disability mentioned by very few pages.

**Conclusions** This study suggests that: (1) public health agencies should increase their web presence in providing correct information on face masks; (2) search engines should improve the information quality criteria in their ranking; (3) the public should be more informed on issues related to the use of masks and disabilities, mental health and stigma arising for those people who cannot wear masks.

## INTRODUCTION

Face masks are widely recommended, along with hygiene and social distancing, as a measure to prevent the spread of COVID-19.<sup>1,2</sup>

## Strengths and limitations of this study

- Three countries (Italy, the USA and the UK) and two languages analysed.
- Sample size of 450 webpages.
- Single search term.
- Information on social networks not analysed.

However, their social acceptance is often problematic with people refusing to use them for several reasons, from health concerns to different social, political and philosophical reasons,<sup>3–5</sup> and even some leaning towards conspiracy theories. A study on the Twitterverse has also identified a stigma about mask-wearing, which can be seen as a mark for COVID-19.<sup>6</sup> On the other hand, there are also contraindications for the use of masks in some health conditions, from respiratory disease to poor mental health or hearing impairment, that are clearly recognised in the WHO guidance.<sup>7</sup> Issues related to face mask availability, particularly in the initial period of the pandemic, also played a role in the hesitancy about their use,<sup>8</sup> with health authorities in some countries initially discouraging the use of surgical masks by the general public.

Health authorities have often changed their advice on the use of masks. For instance, the WHO initially did not recommend their use by the general public on 6 April 2020,<sup>9</sup> but changed the guidance later during the pandemic, on 5 June 2020,<sup>7</sup> with many regional health authorities quickly conforming to the new guidance. There also major differences among different countries, with some governments recommending face masks, some allowing the use of other types of face coverings such as scarves (such as in the UK, where reference is often made to 'face coverings' rather than 'face masks') and some regulating their mandatory use differently in closed spaces, outside, on public

transports etc.<sup>10–12</sup> There has also been major differences in how political leaders in different countries disregarded, supported or mandated the use of masks.<sup>2</sup>

A lack of a worldwide agreement, and sometimes a lack of alignment to the recommendation by the WHO, along with cultural differences, resulted in marked differences in face mask adoption. Data from a survey on 8–14 June carried out by Imperial College London and YouGov (available at: <http://www.coviddatashub.com/>) showed that those stating that they ‘always wore a face mask outside my home’ were 85% in Italy, 56% in the USA and 19% in the UK. Conversely, in each country, 1% in Italy, 11% in the USA and 53% in the UK responded that they would not wear a mask.

People often inform themselves about health-related behaviour on the internet, and there is an abundant literature on the importance of online health information in the phenomenon of vaccine hesitancy and antivax movements.<sup>13–14</sup> The COVID-19 pandemic has caused the development of several conspiracy theories about the origin of the virus, the public health measures to deal with it (eg, face masks, lockdown, isolation, track and tracing apps), as well as the therapeutic strategies against it.<sup>15–16</sup> These are often heavily politicised and polarised, as shown by studies on what has been called an infodemic<sup>17</sup> that includes websites promoting various supplements to ‘boost immunity’, a topic with an important commercial interest.<sup>18</sup>

Although the potential damage that online misinformation regarding the social acceptance of public health measures for COVID-19 has not yet been quantified, two studies have shown that susceptibility to COVID-19-related misinformation can influence adherence to vaccines and non-pharmacological interventions,<sup>19</sup> including face masks.<sup>20</sup> The fact that preventative measures are highly controversial could make the public particularly sensitive to information obtained online. We, therefore, wondered whether different behaviours in different countries could be related to the type of online information on the potentially negative effects of face masks and their association with conspiracy theories. Using a methodology and a workflow previously used to study antivaccine information in different countries,<sup>21–22</sup> we analysed the websites returned by searching the string ‘are face masks dangerous’ in Italy, the UK and the USA. Because different search engines in different countries and languages can differ significantly in the quality of the information they provide on health topics,<sup>21–22</sup> we used three: Bing, Duckduckgo (a privacy-savvy, no-tracking engine) and Google. In fact, as we discussed elsewhere,<sup>22</sup> the algorithm used by Google was updated in 2019 to improve the quality of webpages providing financial or health-related information (‘your money your life’ pages). One of the key criteria to provide a higher ranking to health-related information is that the information provided should be aligned with the scientific consensus. These quality criteria seem not to have been implemented by other search engines that the public may use.<sup>22</sup> Although Google is the most widely used search

engine with a market share of over 90% according to Statcounter.com, its use might have provided a sample of webpages missing those with a particular low quality and that users could find either using other search engines or finding them on social networks.

We analysed the first 50 webpages, each of them returned in the different localisations in English and Italian. Their content was analysed in terms of their overall stance towards masks, the presence of conspiracy information and intrinsic transparency features (such as the presence of author, date, references) and the issues discussed in the context of the use of face masks. Finally, we looked at the impact of published guidance on the information online over time.

## METHODS

### Search strategy

The search string ‘are face masks dangerous’ (with no question mark) was decided as it was among the top suggestions by Google when typing ‘are face mask’. The corresponding search string in Italian was ‘le mascherine sono pericolose’. Three search engines were used (Bing, Duckduckgo and Google) using their localised (language and country) version or settings. The search was performed in August 2020. Searches in English were performed in Brighton, UK, setting English language and the USA or the UK as country. Searches in Italian was performed in Urbino, Italy, setting Italy as the country (thus excluding Switzerland). Cookies and previous browsing history were cleared before each search to minimise customisation effects.

The first 50 URLs for each search engine result page (SERP) were downloaded to a spreadsheet and subsequently visited for the analysis. All the webpages were archived on the internet archive (<https://archive.org/web/>) on 31 August 2020.

### Analysis

Each webpage was first classified for its typology by two raters (see below for inter-rater agreement). The typologies considered were (based on previous studies on several health topics)<sup>21–24</sup>: blogs, commercial, government, health portals, news, no profit, professional, scientific journals, video and social networks. When an article on a webpage had a date, this was noted as well.

Then the content was analysed for the mention of conspiracy theories, debunking information and stance about masks. Stance was assigned by two raters trying to answer the question ‘would a layperson be discouraged from using face masks after visiting this webpage?’. Webpages that did not contain information about face masks were excluded, and the number of excluded webpages for each SERP is reported in the Results section.

Finally, we recorded whether the websites mentioned any negative aspect of face masks and their use. For this, we specifically searched the text of each webpage for words relating to specific issues: efficacy/effectiveness,

hypercapnia/carbon dioxide, respiratory disease/asthma, hypoxia/oxygen (lack of), infection (increased susceptibility or lack of immunity to), use (instructions, exercise, children, immunosuppression, shortage, mental health).

When specific words were searched, this was done automatically. Briefly, the list of URL for each SERP was used to create a text corpus using the corpus analysis software Sketch Engine (sketchengine.eu) by Lexical Computing, Brno-Královo Pole, Czechia.<sup>25</sup> Sketchengine can build a corpus from a list of webpages using WebBootCaT, an online tool for bootstrapping (<https://www.sketchengine.eu/guide/create-a-corpus-from-the-web/>).<sup>26</sup> Because some URLs do not allow access to robots, not all the webpages were analysed. Of the 101 unique webpages in Italian, only a sample of 68 (67%) could be analysed; for the 131 unique URLs in English, 91 (69%) could be analysed.

### Calculation of inter-rater reliability

As one might expect a degree of subjectivity in the assignment of website typology and even more in the assessment of the stance towards masks, whether positive, neutral or negative, the inter-rater agreement was calculated on a sample of 100 URLs from two raters in Italian (AC and TV) and 89 URLs from two raters in English (SM and DG). Any disagreement was resolved by discussion with a third author (PG). Inter-rater reliability was calculated using GraphPad Prism. The Cohen's kappa was used because in this project, there were only two raters and they rated the same sample. Data are reported as kappa and its 95% CI. As proposed by Landis and Koch,<sup>27</sup> we are considering a kappa between 0.41 and 0.60 as a 'moderate' agreement, between 0.61 and 0.80 as a 'substantial' agreement and between 0.81 and 1.00 as an 'almost perfect' one.<sup>27</sup>

### Statistical analysis

The frequency with which an information occurs in two groups of websites is compared using a 2x2 contingency table, followed by a two-tailed Fisher's exact test using GraphPad Prism V.9.1.2 for Windows. A threshold of  $p < 0.05$  is predefined for significant testing and the actual  $p$  value is reported in the results section for each comparison.

### Data sharing

The full list of pages returned in each SERP and their classification, as well as those excluded, are available allow reanalysis in online supplemental file 1.

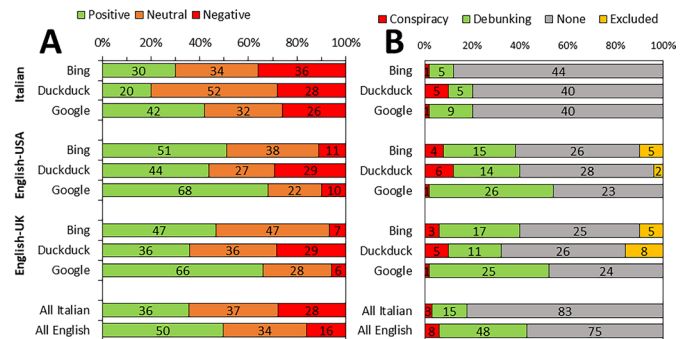
### Patient and public involvement

No patient involved.

## RESULTS

### Inter-rater agreement

The inter-rater agreement for typology in Italian was 90%, with a kappa of 0.742 (95% CI: 0.598 to 0.886; 'substantial' agreement); in English it was 92%, with a kappa of 0.849 (95% CI: 0.743 to 0.955; 'almost perfect' agreement). The inter-rater agreement for the stance on masks in Italian was 79%, with a kappa of 0.686 (95% CI 0.566



**Figure 1** Percentage of webpages in each search engine result page (SERP) with different stances on masks (A) and conspiracy or debunking information (B). (A) Green, positive; orange, neutral; red, negative. All webpages in each language are all the unique webpages in all the SERPs in that language (Italian, 101; English, 131). (B) Conspiracy (red) and debunking (green) information. Number of webpages with conspiracy or debunking information. Total is 50 for each SERP except for Bing USA and Bing UK (n=45), Duckduckgo USA (n=48) and Duckduckgo UK (n=42), as some webpages (in yellow) were excluded for lack of relevance. 'All Italian' and 'all English' represent all the unique webpages in each language (101 and 131, respectively).

to 0.805; 'substantial' agreement); in English it was 66%, with a kappa of 0.441 (95% CI 0.282 to 0.600; 'moderate' agreement).

### Type of webpages returned

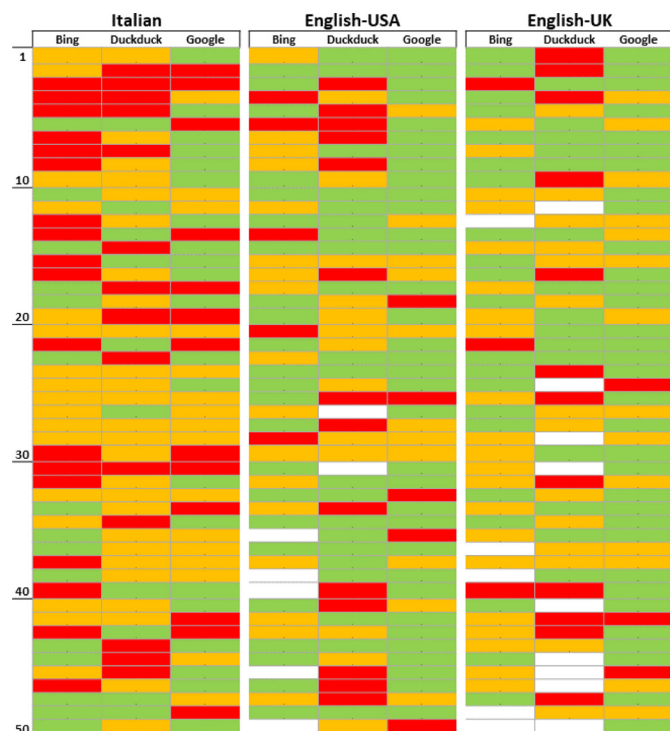
A first analysis looked at the composition of the search results in terms of website typology as described in the Methods section (blogs, commercial, government, health portals, news, no profit, professional, scientific journals, video and social networks). As shown in online supplemental table S1, the majority of websites were news outlets (79% on average in Italian, 64% in English—USA and 69% in English—UK), with few others spread across the different typologies. No more than one government website was returned by each search engine.

### Content analysis

The content of the webpages was assessed in terms of stance about masks and the presence of conspiracy or debunking information. Figure 1 shows the percentage of webpages with a positive, neutral or negative stance towards masks for each search (A) and the presence of conspiracy theories or debunking (B).

It can be seen that the webpages with a positive stance were more frequent in English (65 of 131, 50%) than in Italian (36 of 101, 36%). There was also a difference in the proportion of negative pages in the three search engines, with Google returning the lowest, followed by Bing and Duckduckgo. We then compared all the unique webpages in Italian with all those in English combining the different SERPs and removing duplicates. For each language, the SERPs (three in Italian and six in English) were combined and duplicates removed, resulting in 101 URLs in Italian and 131 in English. We decided to combine





**Figure 2** Overall stance on masks in the first 50 webpages in each search engine result page. Colours indicate the stance: positive (green), neutral (orange) and negative (red). White indicate webpages excluded from the analysis.

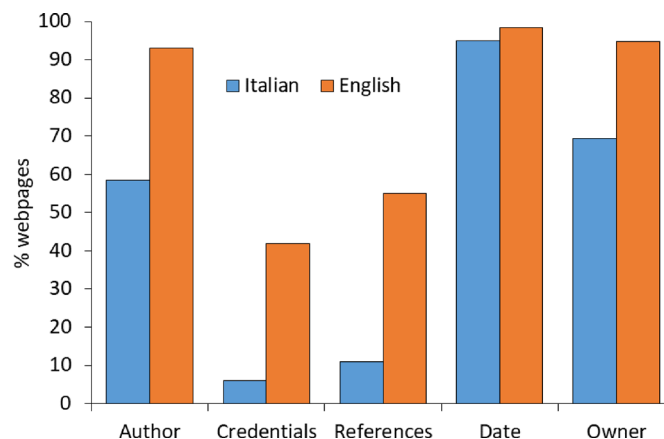
the six SERPs in English because there was a large overlap between the USA and the UK (75 webpages in common, 34 unique for the USA and 22 unique for the UK). As seen from [figure 1A](#), webpages with a negative stance were more frequent in Italian (28 of 101, 28%) than in English (21 of 131, 16%;  $p=0.0354$  by Fisher's exact test). We also analysed whether the webpages-mentioned information related to conspiracy theories or, on the contrary, debunked misinformation. As shown in [figure 1B](#), the highest number of conspiracy webpages was returned by Duckduckgo and the lowest, on average, by Google, that also had the highest number of debunking webpages.

The ranking of the webpages in the different SERPs is reported in [figure 2](#), from which it is clear that mask-negative pages are ranked lower by Google in English but not in Italian.

### Trustworthiness indicators

We then evaluated the transparency/trustworthiness of webpages by looking at the presence of the following criteria: if the author was given and if so their credentials (eg, indicating a degree), external reference to back up the information, the date and the ownership of the website. This was done on the 101 webpages in Italian and the 131 in English, as described above.

The results reported in [figure 3](#) indicate a generally lower presence of trustworthiness indicators in the webpages in Italian. The frequency of each indicator was significantly lower in Italian compared with English with the exception of the presence of a date (Fisher's



**Figure 3** Webpage trustworthiness indicators. Percentage of webpages indicating the presence of transparency indicator (Italian, n=101; English, n=131).

test for each indicator gave the following results: author, credential and references,  $p<0.00001$ ; date,  $p=0.244$ ). This deficiency was particularly evident for the presence of the authors' credentials and of external references. In English, these two indicators were only present in 40%–50% of the webpages analysed.

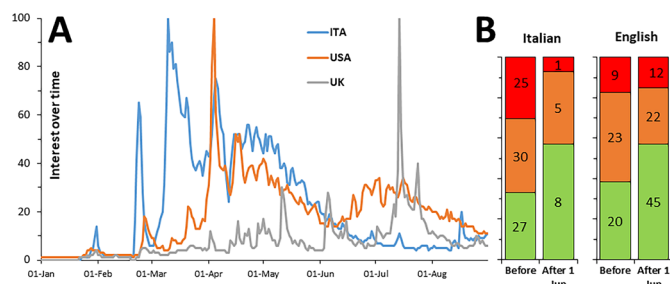
### Variation in the information on masks over time

The guidelines on the use of face masks have changed over time. On 6 April 2020, the WHO recommended their use only by healthcare workers, which on the 5 June was extended to the general public following the publication, on 1 June, of a WHO-sponsored study.<sup>1</sup> In the UK, face masks were made compulsory in public transports on 15 May and this was extended to shops on 24 July. In the USA, regulations were very different among states where they became required in shops or public transports between May and July. On 4 April, the Center for Disease Control (CDC) changed its advice recommending the use of face masks by the general public.<sup>28</sup> In Italy, following the recommendation of the 'Istituto Superiore di Sanità', they were made compulsory in closed spaces on 26 April,<sup>29</sup> although in Lombardy they were required also outside from 11 April.<sup>30</sup>

[Figure 4A](#) shows the frequency of Google searches over the period 1 January to 31 August, using Google trends (<https://trends.google.com/>) in Italy, the USA and the UK.

In Italy (searches in Italian in Italy, excluding Switzerland), there were three peaks in the searches on 23 February, 10 March and 5 April. In the USA, the peaks were on 5 April, and then in the first half of July. In the UK, the peak was on 15 July. The overall trend followed the introduction of the use of face masks, earlier in Italy, then in the USA and later in the UK.

We also analysed the stance about masks over time before and after the 1 of June, when the new WHO recommendations were published. This was done on the 101 unique webpages in Italian and 131 in English. The results, shown in [figure 4B](#), indicate that there was

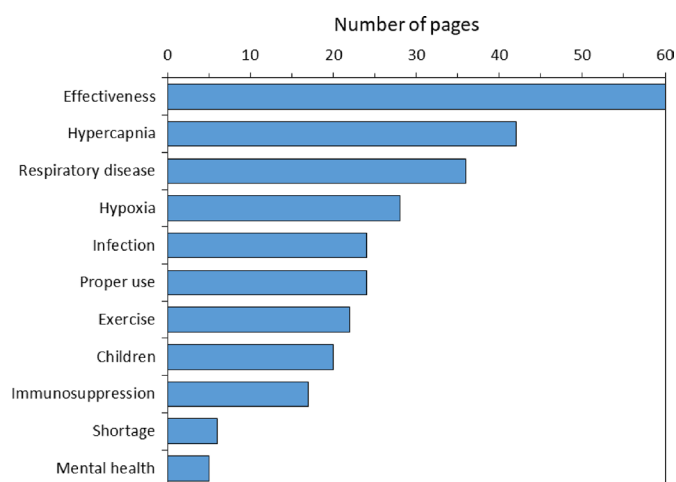


**Figure 4** Impact of public guidelines on online information. (A) Google searches on face masks over time according to Google trends. Data represent search interest relative to the highest point on the chart for the given region and time. A value of 100 is the peak popularity for the term in the time range considered. A score of 0 means there were not enough data. (B) 5. Stance on masks before and after 1 June 2020 in Italian and English (n=101 and 131, respectively). Green, positive; orange, neutral; red, negative. Number of webpages is shown for each language.

a trend for a decrease in negative-stance information in Italian and an increase in positive-stance information in English, although these differences were not statistically significant in either language ( $p=0.103$  in Italian,  $p=0.810$  in English; two-tailed Fisher's exact test).

### Issues associated with the use of masks

We investigated which potential issues associated with the use of masks were mentioned in the 262 unique webpages analysed (Italian and English combined). As shown in figure 5, the issue discussed the most often in total was masks' effectiveness in preventing infection, followed by hypercapnia (due to the accumulation of carbon dioxide), the contraindication in respiratory disease, hypoxia, increased infection (for instance due to reinhalation), difficulties in using masks properly, their use during exercise, the potential contraindications in children, immunosuppression, the limited availability of masks and various effects related to mental health or disabilities.



**Figure 5** Number of webpages mentioning each type of potential issues in all webpages. Total does not add up to 232 because webpages can mention no issues or multiple issues.

When searching specifically for issues related to hearing impairment (searching the words: hearing, deaf, deafness), we found that this issue was mentioned in five of the webpages in English (4%) but only in one of those in Italian (1%). Autism (word) was mentioned only in one webpage in English and in none in Italian. None mentioned stigma that may arise for people for which masks are contraindicated.

### DISCUSSION

A previous study in English and Spanish has analysed the online information on COVID-19 prevention and found that less than half of the website sample provides recommendation on mask-wearing.<sup>31</sup> Our study focused on identifying websites that could potentially promote mask hesitancy.

The present study shows that the majority of the websites returned (60%–80%) were from news outlets, which was not surprising given the newsworthiness of this topic. However, this first level of analysis on the typology of the websites returned by the search engines highlights a lack of websites from government agencies (between 0% and 2% in the different SERPs), and these websites are supposed to provide high-quality information. Professional organisations (universities, hospitals and healthcare organisations), that also usually present high-quality information, represented 12%–14% of the pages returned by Google in English, were much less frequent in the SERPs returned by other search engines in English and nearly absent in Italian (including by Google.it with only 2%). This contrasts with the picture we had in previous studies using the same methodology. For example, in a study on vaccine hesitancy in English and Italian, we found between 2% and 6% government websites and 4%–7% professional websites<sup>21</sup>; in a study on the influenza vaccine in the two languages, government websites returned by Google were between 17% and 42%, and professional ones 4%–19%.<sup>23</sup> This might be due to the smaller effort made by public health agencies in providing information on face masks compared with vaccination campaigns, but also to the higher ranking given by search engines to news sites.

In general, the comparison between the different SERPs confirms the conclusions by our earlier studies<sup>21 22</sup> indicating that Google provides information that is more aligned with the scientific consensus, which we interpret as higher information quality, than other search engines, particularly Duckduckgo that also returned the highest number of webpages reporting conspiracy theories and a low number of debunking webpages. However, Google in Italian provides an overall lower-quality information than in English. Our findings highlight the ethics responsibilities of search engines that act as gatekeepers of online information.

The other aspect analysed is the overall stance of the webpages. In total, there was a higher proportion of pages with a negative stance in Italian than in English and more debunking information in English. Italian webpages also

fared worse in terms of intrinsic transparency indicators. Of note, four of the trustworthiness criteria analysed (presence of author, date, references and ownership) are the four components of the most widely used health information quality score, the *Journal of the American Medical Association (JAMA) Score* for websites.<sup>32</sup>

Although we also considered as an additional indicator of transparency the presence of credentials of the authors (eg, a professional qualification or a university degree), it should be noted that this is not one of the component of the *JAMA Score*. In fact, the credentials of an author are not necessarily a good proxy for information quality as exemplified by the case of Nobel prize Montagnier who, at some point, supported vaccine-hesitant views as well as conspiracy theories on COVID-19.<sup>33 34</sup>

This may seem in contrast with the earlier and higher adoption of face masks in Italy compared with the USA or the UK. It would be inappropriate to draw a causal link as many other factors, particularly cultural ones, are important in the uptake of face masks. However, this suggests that the presence of information depicting masks in a more negative way (associated with less debunking information) does not have a major impact on behaviour of the public.

Recent studies suggested that the quality of online information on COVID-19 can significantly affect adoption of public health recommendation.<sup>35 36</sup> The present study highlights the importance for public health organisations to monitor the issues mentioned online in order to more effectively promote the use of non-pharmacological interventions.

### Strengths and limitations

Because we made some comparison with our previous studies on the online information on vaccines, it is important to note that webpages coded as 'negative' on face masks should not be necessarily viewed as 'anti-science' in the same way we consider those promoting vaccine hesitancy. In fact, the scientific consensus and guidance on masks changed significantly over time and initially most health authorities, including the WHO, discouraged their use by the general public. In this context, the only webpages that could be compared with antivaccine ones are those reporting conspiracy theories but, as shown in figure 1B, there were few of them. We are aware of the well-reported spread of conspiracy theories on COVID-19 and public health measures in the infosphere<sup>15 16</sup> and it is possible that conspiracy theories are more present on social networks. In fact, a recent study analysing over 50 000 tweets against wearing masks in English reported that, on Twitter, conspiracy theories are present in one of four tweets.<sup>37</sup> Another aspect that this research highlighted is the type of issues of face masks that are discussed in websites. Some issues raised frequently are not scientifically based, such as the accumulation of CO<sub>2</sub> and hypercapnia, hypoxia and the increased risk of infection due, for instance, to reinhaling the virus or suppressing the immune system.

However, other issues are less controversial, such as the contraindications in children or people with respiratory complications or psychiatric conditions. Interestingly, few websites mention contraindications in people with some mental health conditions or the consequences of mask use on hearing-impaired persons who lip-read.<sup>7</sup> People that cannot wear masks can be subject to stigma, which could become an increasingly important problem with the increased use of face masks. This problem is particularly evident with issues such as deafness or autism that are mentioned in English webpages but nearly absent in the information in Italian. Of note, a recent study on the Anglophone Twitterverse has not identified any cluster, suggesting that discussion around this is probably a small niche compared with politically motivated antimasker discourse<sup>3</sup> and a recent study highlighted the issue that polarised views can lead to stigma towards both people with and without masks.<sup>38</sup>

Finally, it is important to note the limitations of the study and its methodology. An obvious limitation is the use of a search string that focused on dangers associated with masks. While this was chosen to address the research question of studying information that could promote mask hesitancy, it clearly resulted in returning more mask-negative information that present in the whole web. For instance, users can refer to the dangers posed by masks using the expression 'side effects'. A search of 100 websites on Google for face masks and 'dangerous' versus 'side effects' on 5 May 2021 gave only 2 overlapping results in Italian and 16 in English (data not shown). It could well be that the use of a more technical term, such as 'side effect' as opposed to the more plain-language 'dangerous' would return more scientifically sound webpages, as the use of more technical terms in the search query can return higher-quality information.<sup>39</sup>

Also, a limitation is the use of search engines to obtain a sample of the infosphere. Because search engines aim to return high-information-quality webpages, these could have not returned the lower-quality webpages. On the other hand, there are no means of obtaining a random, non-ranked, sample of websites without using search engines, that all have proprietary, non-disclosed, ranking criteria. Also, the comparison of two different languages that we attempted might be influenced by cultural differences in the social acceptance of masks and different ranking by the search engines. It should not be forgotten that websites are not the only form of information online. Studies have shown that one quarter of YouTube videos spread misinformation on COVID-19. Because search engines return very few videos, we did not have a sufficient number to try to analyse them separately. Studies searching video sources and social networks for information specifically on face masks should also be done in order to have a more comprehensive picture of the information available online. Social networks, in particular, could play a more important role than websites in promoting conspiracy theories through an 'information bubble' effect.



## Policy implications

This study provides some indication to public health authorities, journalists and healthcare professionals, as well as internet companies, to correctly inform the public. In particular:

1. The scarce presence of governmental websites suggests that there is room for improving the overall information quality on this topic in the infosphere by a stronger presence of online information by public health agencies but also with news articles citing websites from public health agencies as sources.
2. Search engines such as Duckduckgo, that rightly focus on privacy and no-tracking, should improve the quality of the health information they return and Google could improve their ranking criteria in Italian.
3. The public should be provided with more information on issues related to the use of masks and disabilities, mental health issues and stigma arising for those who cannot wear masks.

**Twitter** Shaily Mehta @shailu\_x and Pietro Ghezzi @PietroGhezzi

**Contributors** PG and SM designed the research. AC, DG, PG, SM and TV performed the research and analyzed the data. PG wrote the manuscript.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** All data relevant to the study are included in the article or uploaded as supplemental information. All data relevant to the study are included in the article or uploaded as supplemental information as an Excel spreadsheet.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

## ORCID iD

Pietro Ghezzi <http://orcid.org/0000-0003-0911-8358>

## REFERENCES

- 1 Chu DK, Akl EA, Duda S, *et al*. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *Lancet* 2020;395:1973–87.
- 2 Peeples L. Face masks: what the data say. *Nature* 2020;586:186–9.
- 3 Sanders A, White R, Severson L. Unmasking the conversation on masks: natural language processing for topical Sentiment analysis of COVID-19 Twitter discourse. *medRxiv* 2020.
- 4 Bhasin T, Butcher C, Gordon E, *et al*. Does Karen wear a mask? The gendering of COVID-19 masking rhetoric. *Int J Sociol Soc Policy* 2020;40:929–37.
- 5 Palmer CL, Peterson RD. Toxic Mask-ularity: the link between masculine toughness and affective reactions to mask wearing in the COVID-19 era. *Pol Gen* 2020;16:1044–51.
- 6 Li Y, Twersky S, Ignace K, *et al*. Constructing and communicating COVID-19 stigma on Twitter: a content analysis of Tweets during the early stage of the COVID-19 outbreak. *Int J Environ Res Public Health* 2020;17:6847.
- 7 World Health Organization. Advice on the use of masks in the context of COVID-19: interim guidance, 5 June 2020: World Health organization, 2020. Available: [https://www.who.int/publications/i/item/advice-on-the-use-of-masks-in-the-community-during-home-care-and-in-healthcare-settings-in-the-context-of-the-novel-coronavirus-\(2019-ncov\)-outbreak](https://www.who.int/publications/i/item/advice-on-the-use-of-masks-in-the-community-during-home-care-and-in-healthcare-settings-in-the-context-of-the-novel-coronavirus-(2019-ncov)-outbreak) [Accessed 09 Oct 2020].
- 8 Cohen J, Rodgers YvanderM, van der Meulen Rodgers Y. Contributing factors to personal protective equipment shortages during the COVID-19 pandemic. *Prev Med* 2020;141:106263.
- 9 World Health Organization. Advice on the use of masks in the context of COVID-19: interim guidance, 6 April 2020: World Health organization, 2020. Available: <https://apps.who.int/iris/handle/10665/331693> [Accessed 09 Oct 2020].
- 10 Martinelli L, Kopilaš V, Vidmar M, *et al*. Face masks during the COVID-19 pandemic: a simple protection tool with many meanings. *Front Public Health* 2020;8:947.
- 11 Cypionka T, Greenhalgh T, Bassler D, *et al*. Masks and Face Coverings for the Lay Public : A Narrative Update. *Ann Intern Med* 2021;174:511–20.
- 12 Felter C, Bussemaker N. Which countries are requiring face masks. *Council on* 2020.
- 13 Kata A. Anti-vaccine activists, Web 2.0, and the postmodern paradigm—an overview of tactics and tropes used online by the anti-vaccination movement. *Vaccine* 2012;30:3778–89.
- 14 Jolley D, Douglas KM. The effects of anti-vaccine conspiracy theories on vaccination intentions. *PLoS One* 2014;9:e89177.
- 15 Ball P, Maxmen A. The epic battle against coronavirus misinformation and conspiracy theories. *Nature* 2020;581:371–4.
- 16 Biddlestone M, Green R, Douglas KM. Cultural orientation, power, belief in conspiracy theories, and intentions to reduce the spread of COVID-19. *Br J Soc Psychol* 2020;59:663–73.
- 17 Caulfield T. Pseudoscience and COVID-19 - we've had enough already. *Nature* 2020. doi:10.1038/d41586-020-01266-z. [Epub ahead of print: 27 Apr 2020].
- 18 Wagner DN, Marcon AR, Caulfield T. “Immune Boosting” in the time of COVID: selling immunity on Instagram. *Allergy Asthma Clin Immunol* 2020;16:1–5.
- 19 Roozenbeek J, Schneider CR, Dryhurst S, *et al*. Susceptibility to misinformation about COVID-19 around the world. *R Soc Open Sci* 2020;7:201199.
- 20 Hornik R, Kikut A, Jesch E, *et al*. Association of COVID-19 misinformation with face mask wearing and social distancing in a nationally representative US sample. *Health Commun* 2021;36:6–14.
- 21 Arif N, Al-Jefri M, Bizzi IH, *et al*. Fake news or weak science? visibility and characterization of Antivaccine Webpages returned by Google in different languages and countries. *Front Immunol* 2018;9:1215.
- 22 Ghezzi P, Bannister PG, Casino G, *et al*. Online information of vaccines: information quality, not only privacy, is an ethical responsibility of search Engines. *Front Med* 2020;7:400.
- 23 Maki A, Evans R, Ghezzi P. Bad news: analysis of the quality of information on influenza prevention returned by Google in English and Italian. *Front Immunol* 2015;6:616.
- 24 Cassa Macedo A, Oliveira Vilela de Faria A, Ghezzi P. Boosting the immune system, from science to myth: analysis of the Infosphere with Google. *Front Med* 2019;6:165.
- 25 Kilgariff A, Baisa V, Bušta J, *et al*. The sketch engine: ten years on. *Lexicography* 2014;1:7–36.
- 26 Baroni M, Kilgariff A, Pomikálek J, *et al*. WebBootCaT: a web tool for instant corpora. *Proceeding of the EuraLex Conference*, 2006.
- 27 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
- 28 Center for Disease Control. How to Protect Yourself & Others, 2020. Available: [https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html?CDC\\_AA\\_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fprepare%2Fprevention.html](https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fprepare%2Fprevention.html), archived at [https://web.archive.org/web/2021008145107/https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html?CDC\\_AA\\_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fprepare%2Fprevention.htm](https://web.archive.org/web/2021008145107/https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fprepare%2Fprevention.htm) [Accessed 09 Oct 2020].
- 29 Presidente del Consiglio dei Ministri. Ulteriori disposizioni attuative del decreto-legge 23 febbraio 2020, n. 6, recante misure urgenti in materia di contenimento e gestione dell'emergenza epidemiologica



- da COVID-19, applicabili sull'intero territorio nazionale. *Gazzetta Ufficiale della Repubblica Italiana*;108 del 27-04-2020.
- 30 Lombardia R. Ulteriori misure per La prevenzione E gestione dell'emergenza epidemiologica dA COVID-19, 2020. Available: <https://www.regione.lombardia.it/wps/wcm/connect/e22eba18-e846-4445-a6ff-9cad927a8569/Ordinanza+528.pdf?MOD=AJPERES> [Accessed 8 May 2021].
  - 31 Hernández-García I, Giménez-Júlvez T. Assessment of health information about COVID-19 prevention on the Internet: infodemiological study. *JMIR Public Health Surveill* 2020;6:e18717.
  - 32 Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewor--Let the reader and viewer beware. *JAMA* 1997;277:1244–5.
  - 33 Salzberg S. *Nobel laureate joins anti-vaccination crowd at Autism One*. Forbes, 2012.
  - 34 France-Presse A. *COVID-19: la théorie d'un virus fabriqué vivement contestée*, 2020.
  - 35 Bridgman A, Merkley E, Loewen PJ, *et al*. The causes and consequences of COVID-19 misperceptions: understanding the role of news and social media. *Harv Kennedy Sch Misinformation Rev* 2020;1.
  - 36 Barua Z, Barua S, Aktar S, *et al*. Effects of misinformation on COVID-19 individual responses and recommendations for resilience of disastrous consequences of misinformation. *Prog Disaster Sci* 2020;8:100119.
  - 37 Al-Ramahi M, Elnoshokaty A, El-Gayar O, *et al*. Public discourse against masks in the COVID-19 era: Infodemiology study of Twitter data. *JMIR Public Health Surveill* 2021;7:e26780.
  - 38 Betsch C, Korn L, Sprengholz P, *et al*. Social and behavioral consequences of mask policies during the COVID-19 pandemic. *Proc Natl Acad Sci U S A* 2020;117:21851–3.
  - 39 Dy CJ, Taylor SA, Patel RM, *et al*. The effect of search term on the quality and accuracy of online information regarding distal radius fractures. *J Hand Surg Am* 2012;37:1881–7.